Cell

Structural Basis for Promoter −10 Element Recognition by the Bacterial RNA Polymerase σ Subunit

Andrey Feklistov[1,*] and Seth A. Darst[1,*]
[1]The Rockefeller University, 1230 York Avenue, New York, NY 10065, USA
*Correspondence: afeklistov@rockefeller.edu (A.F.), darst@rockefeller.edu (S.A.D.)
DOI 10.1016/j.cell.2011.10.041

SUMMARY

The key step in bacterial promoter opening is recognition of the −10 promoter element ($T_{-12}A_{-11}T_{-10}A_{-9}A_{-8}T_{-7}$ consensus sequence) by the RNA polymerase σ subunit. We determined crystal structures of σ domain 2 bound to single-stranded DNA bearing −10 element sequences. Extensive interactions occur between the protein and the DNA backbone of every −10 element nucleotide. Base-specific interactions occur primarily with A_{-11} and T_{-7}, which are flipped out of the single-stranded DNA base stack and buried deep in protein pockets. The structures, along with biochemical data, support a model where the recognition of the −10 element sequence drives initial promoter opening as the bases of the nontemplate strand are extruded from the DNA double-helix and captured by σ. These results provide a detailed structural basis for the critical roles of A_{-11} and T_{-7} in promoter melting and reveal important insights into the initiation of transcription bubble formation.

INTRODUCTION

Transcription initiation is a major point for the regulation of gene expression, and the DNA-dependent RNA polymerase (RNAP) is the central enzyme of transcription. In bacteria, the promoter-specificity σ factor combines with the core RNAP to form the holoenzyme, which carries out all steps of initiation (Murakami and Darst, 2003).

The σ factor recruits core RNAP to sites of transcription initiation through recognition of specific DNA sequences called promoters (Shultzaberger et al., 2007). The group 1, or primary, σ factors (σ^{70} in *Escherichia coli*, σ^{A} in *Thermus aquaticus*) are responsible for the bulk of transcription during log-phase growth and are essential for viability.

The −10 element (or Pribnow box) is the most highly conserved and essential bacterial promoter motif (Figures 1A and 1B) (Hook-Barnard and Hinton, 2007; Shultzaberger et al., 2007). Once bound to the promoter in a closed (double-stranded) complex (RP_c), σ^{A}-holoenzyme spontaneously isomerizes to the transcription-competent open complex (RP_o), in which the DNA from within the −10 element downstream to the transcription start site (TSS, +1) is strand separated to form the transcription bubble (deHaseth et al., 1998; the numbering scheme for the transcription register, with the transcription start site as +1 and negative and positive numbers corresponding to upstream and downstream positions, respectively, is denoted in Figure 1A).

Based on the observation that σ-conserved region 2 (Lonetto et al., 1992) contains a number of invariant aromatic and basic residues (Figure 1C), Helmann and Chamberlin (1988) proposed that a function of σ is to bind one of the DNA strands within the nascent transcription bubble to stabilize the strand-separated state. Subsequent studies have shown that the invariant aromatic and basic residues of σ play key roles in −10 element recognition and transcription bubble formation (deHaseth and Helmann, 1995). Alanine substitutions of many of the invariant aromatic residues result in promoter-melting defects (Juang and Helmann, 1994). Holoenzyme sequence-specifically binds single-stranded DNA (ssDNA) corresponding to the −10 element nontemplate-strand (nt-strand) sequence (Marr and Roberts, 1997; Roberts and Roberts, 1996; Savinkova et al., 1988), and this activity has been mapped to structural domain 2 of σ (σ_2; Severinova et al., 1996; Young et al., 2001). Free σ, or fragments containing σ_2, have sequence-specific ssDNA-binding activity for the nt-strand −10 element sequence (Feklistov et al., 2006; Sevostyanova et al., 2007; Zenkin et al., 2007). In fact, the essence of RNAP promoter-melting activity is localized to nt-strand −10 element/σ_2 contacts (Young et al., 2004). These findings support a model whereby σ_2-mediated capture of nt-strand bases of the −10 element extruded from the DNA double-helix underlies the initiation of strand separation and provides crucial stability to RP_o. In addition to its role in initiation, sequence-specific contact between σ_2 and ssDNA can regulate transcript elongation by inducing a pause when σ_2 recognizes −10-element-like sequences in the nt-strand of the transcription bubble (Ring et al., 1996).

The structure of *T. aquaticus* (*Taq*) σ^{A}-holoenzyme with a fork-junction promoter fragment revealed that the invariant aromatic residues of σ^{A}_2 are perfectly positioned to interact with exposed bases of the −10 element nt-strand, but details were not revealed at 6.5 Å resolution (Murakami et al., 2002). To provide a high-resolution structural description of the key protein/DNA interaction in RP_o formation, we determined X-ray crystal structures of a *Taq* σ^{A} fragment comprising structural domains

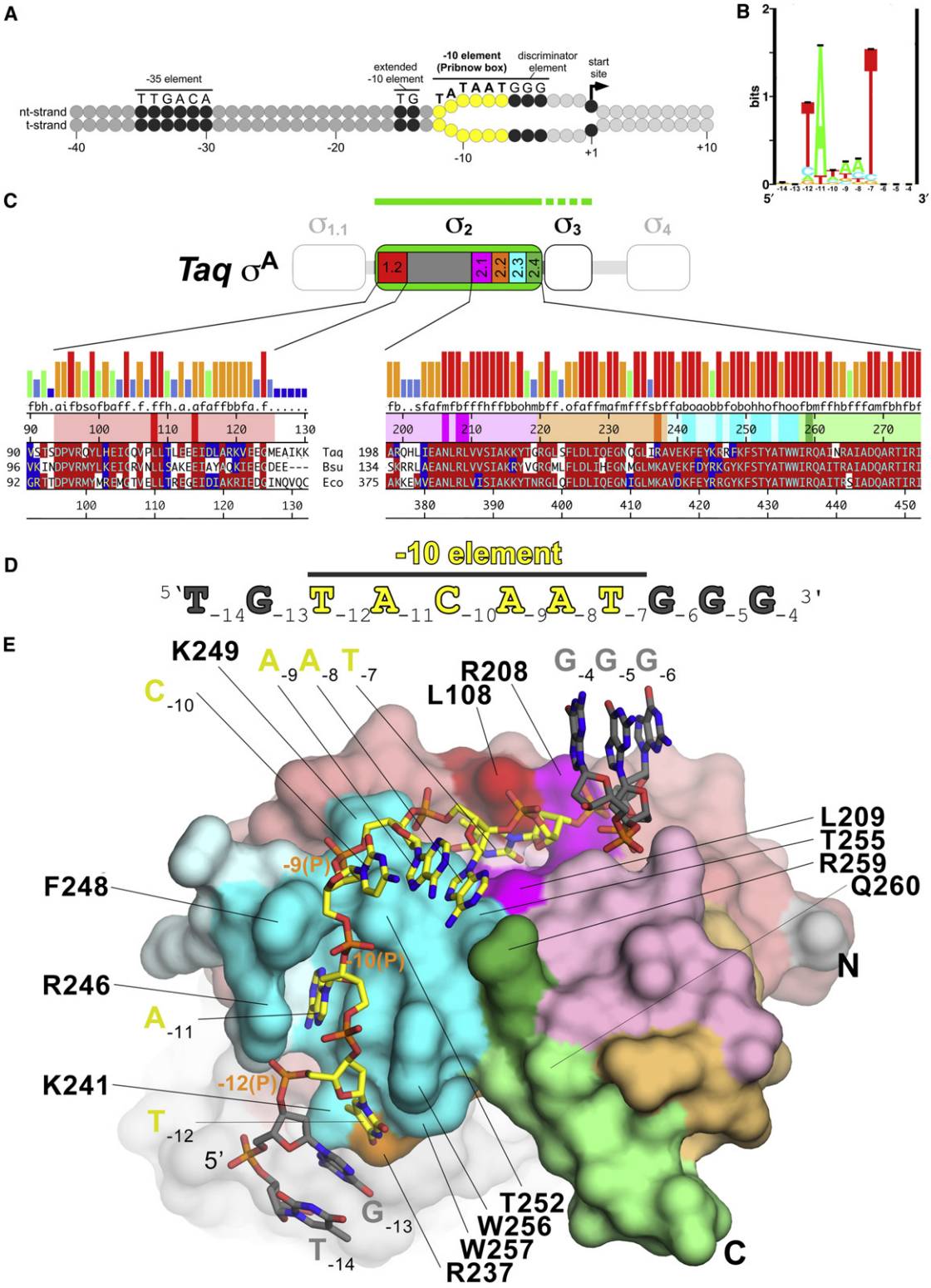

Figure 1. Bacterial Promoter Architecture, −10 Element, RNAP σ Factor, Crystallization Oligonucleotide, and Structural Overview; See also Figure S1 and Table S1

(A) Promoter motifs recognized by primary bacterial RNAP σ factors. Gray circles represent the DNA nucleotides (top, nt-strand; bottom, t-strand). The extent of the transcription bubble in RP_o is illustrated (separated circles). Promoter motifs recognized by primary σ factors are colored black (−35 element,

2 and 3 (σ^A_{2-3}; Campbell et al., 2002) bound to ssDNA containing the −10 element nt-strand sequence (Figures 1C–1E).

RESULTS

Overall Structure of the Single-Stranded −10 Element/σ_2 Complex

The oligonucleotide sequence used for cocrystallization (Figure 1D) is based on the conserved motif discovered in ssDNA aptamers selected for binding to free *Taq* σ^A in vitro (Feklistov et al., 2006). The ssDNA oligonucleotides bound *Taq* σ^A_{2-3} (Figure 1C) with a dissociation constant (K_D) more than three orders of magnitude lower than a control, anticonsensus sequence (Figure S1A available online). The most detailed model (with the $\mathbf{T}_{-12}\mathbf{A}_{-11}\mathbf{C}_{-10}\mathbf{A}_{-9}\mathbf{A}_{-8}\mathbf{T}_{-7}$ −10 element; Figures 1D and 1E) was refined at 2.1 Å resolution to an R/R_{free} of 0.194/0.237 (Table S1; Figure S1B). The structure with the $\mathbf{T}_{-12}\mathbf{A}_{-11}\mathbf{T}_{-10}\mathbf{A}_{-9}\mathbf{A}_{-8}\mathbf{T}_{-7}$ −10 element was essentially identical (root-mean-square deviation of 0.165 Å over all atoms). Although biochemical analysis of the crystal contents established that σ^A_3 was present in the crystals, electron density for the domain was absent, and it was presumed disordered.

The ssDNA is draped across a highly conserved, positively charged surface of σ^A_2 (Figures 1E and S1C), with a 90° turn in the DNA backbone between the −11 and −10 positions (Figure 1E). Extensive interactions occur between the protein and the DNA backbone of every nucleotide from −12 to −6 (Figure 2A). Base-specific interactions occur primarily with $\mathbf{T}_{-12}$, $\mathbf{A}_{-11}$, $\mathbf{A}_{-8}$, and $\mathbf{T}_{-7}$, especially $\mathbf{A}_{-11}$ and $\mathbf{T}_{-7}$, which are notably flipped out of the ssDNA base stack and entirely buried in protein pockets (Figure 1E).

The ssDNA/protein complex results in the burial of 1,096 Å^2 total molecular surface area. On the DNA, the interactions occur almost entirely within the −10 element (92% of the buried surface area). The DNA interacts with residues from all of the σ-conserved regions (1.2, 2.1, 2.2, 2.3, and 2.4; Figures 1C and 1E) (Lonetto et al., 1992), with the bulk of the interactions occurring within σ regions 2.3 (73% of the buried surface area) and 2.1 (23%).

Based on previous studies, we expected to observe interactions between the downstream discriminator element ($G_{-6}G_{-5}G_{-4}$) and residues of σ^A region 1.2 (Feklistov et al., 2006; Haugen et al., 2006, 2008). In fact, the bases of the discriminator element do not interact with the protein but peel away and participate in crystal contacts with GGG motifs from symmetry-related complexes, forming an unexpected G-quadruplex structure that is unlikely to be relevant to σ factor function but plays a critical role in crystal packing (Figure S1D). Nearby, the σ^A structure features a shallow, positively charged channel that likely accommodates the GGG motif in the physiological complex (Figure S1C) (Haugen et al., 2008).

Recognition of T_{-12} Consistent with Base-Paired −12 Position in RP_o

T is strongly favored at −12, the upstream position of the −10 element (Figure 1B). In general, promoter mutations that substitute $\mathbf{T}_{-12}$ with another base weaken promoter activity/binding, and mutations that substitute another base with T strengthen promoter activity/binding (Moyle et al., 1991). Optimal binding of RNAP holoenzyme to fork-junction promoter probes, which have a mostly single-stranded −10 element, required the base pair at −12 (Guo and Gralla, 1998), suggesting that the −12 position normally remains base-paired even in RP_o (Figure 1A). Nevertheless, $\mathbf{T}_{-12}$ can also be recognized in the context of ssDNA (Feklistov et al., 2006; Roberts and Roberts, 1996; Sevostyanova et al., 2007).

The three 5′ nucleotides of the ssDNA ($^{5'}T_{-14}G_{-13}\mathbf{T}_{-12}$) hang off the edge of the protein structure, with only $\mathbf{T}_{-12}$ making significant interactions with the protein (Figures 1E and 2). The three nucleotides maintain a structure similar to one strand of a B-form double-helix, except that the base of G_{-13} is in the *syn* conformation. Modeling in a B-form double-helix reveals that there is ample space for the paired t-strand (−14 to −12), but continuation of the double-helix downstream to −11 is blocked by the invariant W-dyad (W256/W257) and other elements of the protein (Figures 2B and 3A).

$\mathbf{T}_{-12}$ is propped against the W-dyad, with W256 making extensive van der Waals interactions primarily with the $\mathbf{T}_{-12}$

extended −10 element, discriminator element) or yellow (−10 element or Pribnow box), with the consensus sequences above. The position with respect to the transcription start-site (+1) is denoted below.

(B) Sequence logo (Schneider and Stephens, 1990) for the bacterial primary σ factor −10 element (adapted from Shultzaberger et al., 2007).

(C) *Taq* σ^A domain architecture, crystallization construct, and sequence characteristics. The domain architecture of *Taq* σ^A is represented schematically (structured domains, thick regions; flexible linker, thin regions; Campbell et al., 2002). The crystallization construct, comprising σ_{2-3}, is highlighted. The green bar above indicates that σ_2 was ordered in the crystal structure (solid bar), whereas σ_3 was disordered (dashed bar). The conserved regions within σ_2 (Lonetto et al., 1992) are labeled and color-coded. Expanded below is a sequence alignment of the σ_2 conserved regions for *Taq* σ^A, *Bacillus subtilis* (*Bsu*) σ^A, and *E. coli* σ^{70}. Sequences are shown in one-letter amino acid code. Numbers at the beginning of each line indicate the amino acid positions. Number scales at the top and bottom indicate amino acid position in *Taq* σ^A, and *E. coli* σ^{70}, respectively. The sequence blocks are color-coded according to the schematic above; the darker bands of color denote protein side chains that interact with the DNA. Amino acid identity in >50% of the sequences is indicated by a red background, amino acid similarity by a blue background. Groups of residues considered similar are the following: ST (h), RK (b), DE (a), NQ (m), FYW (o), and ILVM (f). The histogram at the top represents the level of sequence conservation (using the groups denoted above) at each position in an alignment of 50 primary bacterial RNAP σ factors (Campbell et al., 2002). Sequence conservation of 100% is represented by a tall red bar, less than 20% by a small dark blue bar, and intermediate levels are represented by orange, light green, and light blue bars.

(D) Crystallization oligonucleotide. Synthetic 11-mer ssDNA oligonucleotide used for crystallization (Feklistov et al., 2006). The −10 element is colored yellow and labeled.

(E) Structural overview. The σ^A_2 protein is shown as a molecular surface and color-coded as in (C) (conserved regions shown in pale colors, and residues with side chains that interact with the DNA shown in darker colors). Selected residues are labeled (see text). The ssDNA is shown with carbon atoms color-coded as in (D). Other atoms are colored as follows: nitrogen, blue; oxygen, red; phosphorous, orange.

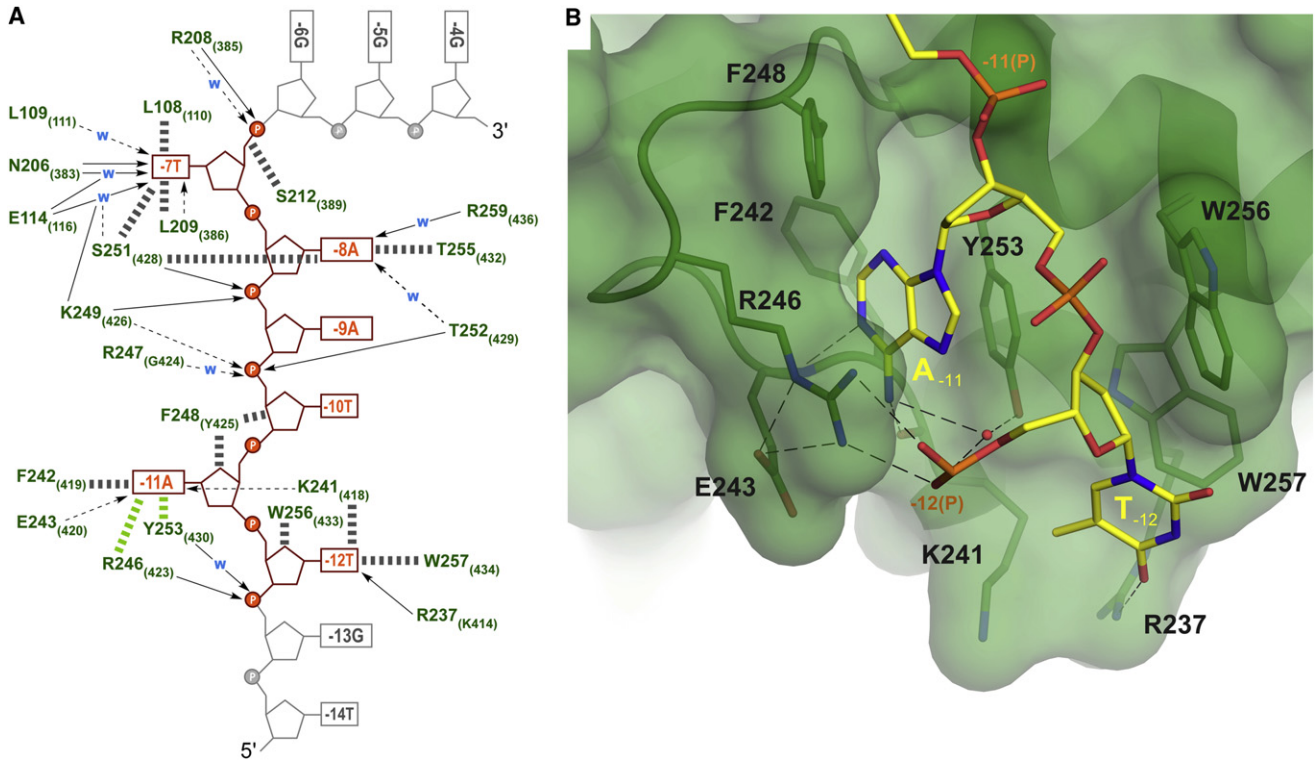

Figure 2. DNA/Protein Contacts and Recognition of $T_{-12}A_{-11}$

(A) Schematic showing DNA/protein contacts. Arrows indicate direct or water-mediated (blue W) hydrogen bonds from side chain (solid) or main chain (dashed) protein atoms. Van der Waals contacts are shown by thick gray dashed lines; stacking/cation-π interactions are shown by thick green dashed lines. The numbers indicate amino acid positions for *Taq* σ^A and *E. coli* σ^{70} (in brackets). Gray areas of the oligo indicate regions without biologically relevant contacts.

(B) The $\sigma^A{}_2$ protein is shown as an α-carbon backbone ribbon and with a transparent molecular surface (green). Selected protein side chains are shown and labeled. The ssDNA is shown as in Figure 1E (nucleotides upstream of $\mathbf{T}_{-12}$ have been omitted for clarity). An ordered water molecule is shown as a small red sphere. Polar interactions (H-bonds and/or salt bridges) are denoted (gray dashed lines).

deoxyribose moiety (as well as with the base), and W257 making an "edge-on" van der Waals interaction with the $\mathbf{T}_{-12}$ pyrimidine ring (Figure 2B). The $\mathbf{T}_{-12}$ 5′-phosphate [−12(P)] is held by polar interactions with R246. These interactions would likely occur regardless of the identity of the −12 base. Explaining the preference for T at this position, R237, reaches over from the σ region 2.2 α helix and forms a hydrogen bond (H-bond) with the O4 atom of $\mathbf{T}_{-12}$ [$\mathbf{T}_{-12}$(O4)], and the aliphatic side chain of K241 makes van der Waals contact with the $\mathbf{T}_{-12}$(C5-methyl) (Figure 2B).

The primary role of σ_2 in −10 element recognition was first uncovered in genetic screens for σ mutants that suppressed single-base substitutions in the −10 element. Specifically, it was shown that substitutions at the position corresponding to *E. coli* σ^{70} Q437 (*Taq* σ^A Q260), which is absolutely conserved among group 1 σ's (Campbell et al., 2002; Gruber and Bryant, 1997) (Figure 1C), to H or R allow efficient transcription from mutant promoters having a T to C substitution at position −12 (Kenney et al., 1989; Waldburger et al., 1990).

In the *Taq* $\sigma^A{}_2$/−10 element DNA structure, electron density maps show clear evidence for two, roughly equally populated conformations of Q260. In either conformation, however, the shortest distance between any atom of Q260 and $\mathbf{T}_{-12}$ is 6.3 Å—too far for the genetic results to be explained by a direct interaction between Q260 and the $\mathbf{T}_{-12}$ base (Figure 1E). However, modeling of the t-strand A base-paired to $\mathbf{T}_{-12}$ places the major-groove edge of this base within H-bonding distance of Q260 (Figure 3A). Surveys of protein/DNA interactions (Hoffman et al., 2004; Luscombe et al., 2001) point to a strong preference for Q to interact with the major-groove edge of A, whereas H and R strongly prefer G (Figure 3A), which corroborates our modeling and explains previous genetic results.

Q260 and other σ region 2.4 residues implicated in −10 element recognition lie on a long α helix that is roughly perpendicular to the trajectory of the promoter DNA double-helix (Murakami et al., 2002). Because of this, amino acid side chains of σ do not appear to be able to establish sequence-specific interactions with the double-stranded −10 element (as in RP_c) due to the depth of the major groove. Structural modeling (Figure 3A) suggests that Q260 can recognize the major-groove edge of the −12 bp only when the −11 position and downstream are strand-separated, allowing the −12 position of the resulting upstream fork junction to move closer to the σ region 2.4 α helix.

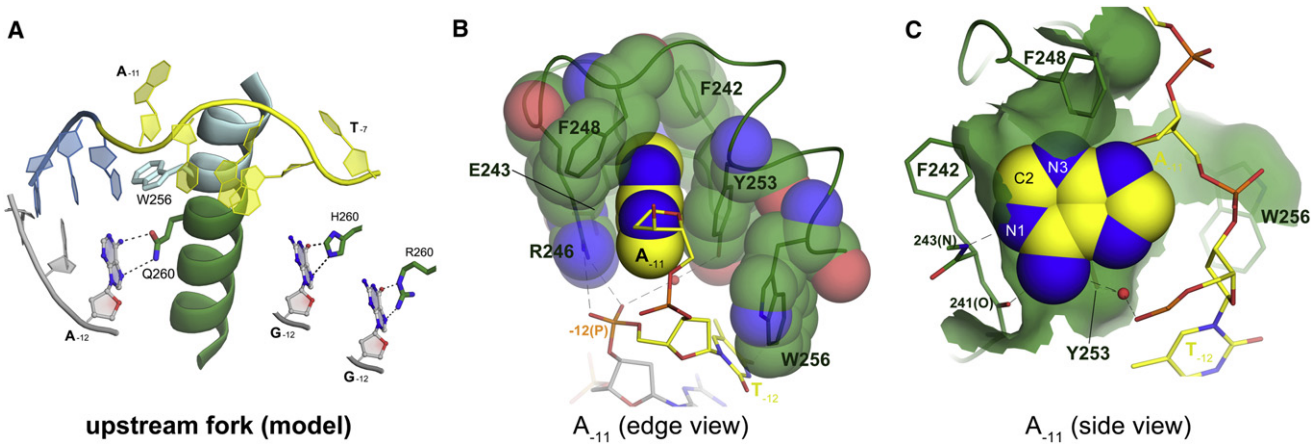

Figure 3. Details of $T_{-12}A_{-11}$ Recognition; See also Figure S2

(A) Modeling of the upstream fork junction and recognition of the $(T/A)_{-12}$ bp. The upstream dsDNA is modeled as a B-form double-helix based on Murakami et al. (2002). In the modeled portion of the DNA, the nt-strand is colored blue; the t-strand, gray. The α helix formed by σ regions 2.3 (turquoise) and 2.4 (green) is shown as a ribbon. The side chains of W256 (blocks extension of DNA double-helix downstream past the −12 bp and occupies the space vacated by the flipped $\mathbf{A}_{-11}$) and Q260 (positioned to interact with the major-groove surface of the t-strand A base-paired to $\mathbf{T}_{-12}$) are shown. Modeling of the recognition of the mutant $(C/G)_{-12}$ base pair by Q260H (Waldburger et al., 1990) or Q260R (Kenney et al., 1989) is shown on the right.

(B) Edge view of $\mathbf{A}_{-11}$ base. The ssDNA is shown with the atoms of the $\mathbf{A}_{-11}$ base in CPK format. An ordered water molecule is shown as a small red sphere. Polar interactions (H-bonds and/or salt bridges) are denoted (gray dashed lines). The $\sigma^{A}{}_{2}$ protein is shown as a worm. Side chains that make up the $\mathbf{A}_{-11}$ binding pocket are shown with transparent CPK atoms.

(C) Side view of $\mathbf{A}_{-11}$ base. The $\mathbf{A}_{-11}$ binding pocket has been sliced near the level of the base and is viewed from the side. The molecular surface of $\sigma^{A}{}_{2}$ within 4.5 Å of the $\mathbf{A}_{-11}$ nucleotide is shown as a transparent surface.

$\mathbf{A}_{-11}$ Is Flipped out of the DNA Base Stack and Buried in a Complementary Hydrophobic Pocket of σ_2

The most highly conserved position of the −10 element is $\mathbf{A}_{-11}$ (Figure 1B). Only a few percent of σ^{A} promoters have a base other than A at this position. Mutations from the consensus A (1) often completely inactivate promoters (Lee et al., 2004; Lim et al., 2001), (2) cause severe defects in RP_o stability (Fenton and Gralla, 2001, 2003a, 2003b; Guo and Gralla, 1998; Lim et al., 2001; Matlock and Heyduk, 2000; Schroeder et al., 2007), and (3) cause defects in binding to nt-strand ssDNA oligonucleotides (Roberts and Roberts, 1996) (Figure S1A).

In addition to interacting with the $\mathbf{T}_{-12}$ backbone, the first W of the invariant W-dyad, W256, also interacts with the $\mathbf{A}_{-11}$ backbone and occupies the space where the $\mathbf{A}_{-11}$ ribose moiety would be if the DNA double-helix extended downstream from −12 (Figures 1E, 2B, and 3A). This necessitates a flip of the entire $\mathbf{A}_{-11}$ nucleotide, which, in turn, removes the $\mathbf{A}_{-11}$ base from the upstream base stack formed by $T_{-14}G_{-13}\mathbf{T}_{-12}$ (Figure 1E). Instead, the $\mathbf{A}_{-11}$ base is completely buried in a hydrophobic protein pocket (Figures 2, 3B, and 3C) (Tsujikawa et al., 2002).

The $\mathbf{A}_{-11}$ pocket is perfectly shaped to fit an A and would poorly accommodate any other base, explaining the high conservation of $\mathbf{A}_{-11}$ in the −10 element (Figure 1B) and the severe effect of promoter mutations at this position. On one face of the $\mathbf{A}_{-11}$ base (the front, or downstream face), Y253 makes a π stack, as predicted by Schroeder et al. (2007). On the opposite face, R246 stacks on the base, forming a cation-π interaction (Wintjens et al., 2000) (Figures 1E, 2B, and 3B). The position of the R246 side chain is stabilized through a polar interaction with the −12(P) (Figures 2 and 3B). In the absence of DNA, the R246 side chain is free to swing into an open configuration, allowing the $\mathbf{A}_{-11}$ base to slip into the pocket.

Studies using nucleotide analogs in place of $\mathbf{A}_{-11}$ revealed a strict requirement for a purine base with no side groups at the N1 and C2 positions (Lee et al., 2004; Matlock and Heyduk, 2000). Furthermore, methylation of $\mathbf{A}_{-11}$ at N3 interfered with holoenzyme binding (Johnsrud, 1978; Siebenlist et al., 1980). In the structure, the back wall of the $\mathbf{A}_{-11}$ pocket forms a tight steric fit with the base that is only possible if the N1, C2, and N3 positions are unsubstituted (Figure 3C). F242 makes a hydrophobic contact with $\mathbf{A}_{-11}$(C2), whereas the polypeptide backbone between residues 241 and 243 makes several H-bonds with the $\mathbf{A}_{-11}$ base. The $\mathbf{A}_{-11}$ pocket is topped by F248, which makes van der Waals contact with $\mathbf{A}_{-11}$(N3) (Figures 3B and 3C). The side chain of E243 does not contact the $\mathbf{A}_{-11}$ base but appears to play an important role by forming the bottom part of the pocket and by making polar interactions with R246 to help stabilize its position (Figure 2B).

$(C/T)_{-10}A_{-9}A_{-8}$ Interact with σ_2 Primarily through the DNA Backbone

The path of the DNA backbone wraps around the surface of the protein with a 90° turn between the −10(P) and −9(P) (Figure 1E). T252 interacts with −9(P), and serves as a fulcrum of the DNA backbone turn (Figures 1E and 4). In this way, T252 plays a critical role. Of all the highly conserved residues of σ region 2.3 (Figure 1C), T252 is the least tolerant to substitution (Waldburger and Susskind, 1994). Changes at this position yield the most severe promoter-melting defects in vitro (Schroeder et al., 2008), and the only substitution that yields functional σ in vivo

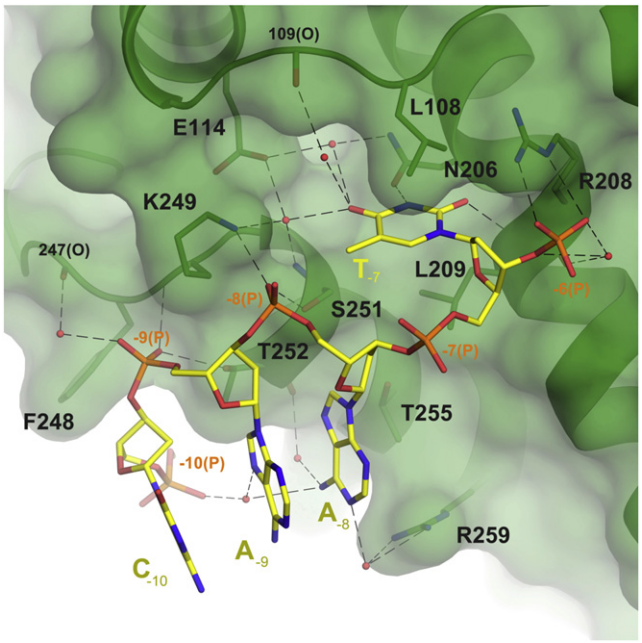

Figure 4. Recognition of $C_{-10}A_{-9}A_{-8}T_{-7}$; See also Figure S2
The σ^{A}_{2} protein, selected protein side chains, the ssDNA, ordered water molecules, as well as polar interactions (H-bonds and/or salt bridges) are shown as in Figure 2B.

is highly conservative S (Waldburger and Susskind, 1994). Comparison of promoter-binding versus -melting activities of the *E. coli* σ^{70} T429A substitution (corresponding to *Taq* σ^{A} T252A) suggests that T252 exerts its critical role at the strand-separation step, after formation of RP_c (Schroeder et al., 2008), consistent with the structure.

After $\mathbf{A}_{-11}$, the most highly conserved position of the −10 element, the next three nucleotides, $\mathbf{T}_{-10}\mathbf{A}_{-9}\mathbf{A}_{-8}$, are the least conserved (Figure 1B), and promoter mutations at these positions generally have less effect on promoter activity than mutations at the −12, −11, or −7 positions. In line with these observations, the −10/−9/−8 nucleotides are primarily bound through extensive interactions with the DNA sugar-phosphate backbone (Figures 2A and 4). The three bases are stacked together and point away from the protein; only the base of $\mathbf{A}_{-8}$ makes van der Waals contact with T255 and also water-mediated contacts to the R259 side chain and T252 main chain (Figure 4).

The −9(P) and −8(P) make extensive polar contacts with protein side chains and main chain, whereas the −7(P) does not interact with the protein (Figures 4 and S2). This explains chemical probing results, which found that ethylation of the −9(P) or −8(P), but not other phosphates in the −10 element, interfered with promoter binding by RNAP (Johnsrud, 1978; Siebenlist et al., 1980).

T_{-7} Is Flipped out of the DNA Base Stack and Buried in a Hydrophilic Pocket of σ_2

The downstream position of the −10 element, $\mathbf{T}_{-7}$, is almost as highly conserved as $\mathbf{A}_{-11}$ (Figure 1B). Promoter mutations at this position also generally have severe consequences for promoter activity (Moyle et al., 1991). The base stack of $\mathbf{C}_{-10}\mathbf{A}_{-9}\mathbf{A}_{-8}$ is prevented from continuing in the downstream direction by T255 and R259 (Figure 4). Instead, the entire $\mathbf{T}_{-7}$ nucleotide is flipped out of the base stack (as predicted by Schneider, 2001) and buried in another protein pocket formed by residues from conserved regions 1.2, 2.1, and 2.3 of σ (Figure 4). Unlike the $\mathbf{A}_{-11}$ protein pocket, the $\mathbf{T}_{-7}$ pocket is (1) spacious compared with the size of the base and (2) hydrophilic in nature. The $\mathbf{T}_{-7}$ pocket accommodates well-ordered water molecules that participate in the recognition of the base (Figures 2A and 4). Every potential interacting moiety of the $\mathbf{T}_{-7}$ base is recognized by the protein.

Although the $\mathbf{T}_{-7}$ pocket is relatively spacious compared to the pyrimidine base, purine bases cannot be accommodated in the pocket. The spatial arrangement of H-bond donors and acceptors of a pyrimidine C_{-7} are not compatible with the $\mathbf{T}_{-7}$ pocket, and a favorable hydrophobic van der Waals interaction would be lost due to the absence of the C5-methyl (Figure S2A).

A salt bridge between R208 and the −6(P) is the final biologically relevant DNA/protein contact (Figure 4). Downstream, $G_{-6}G_{-5}G_{-4}$ turn away from the protein to form the intermolecular G-quartet structure that participates in crystal packing (Figures 1E, S1C, and S1D).

Recognition of the −10 Element Sequence Is Coupled with Strand Separation

It has been established that the σ_2 sequence specifically recognizes the nt-strand of the −10 element in RP_o (Marr and Roberts, 1997; Roberts and Roberts, 1996; Savinkova et al., 1988), mediated by universally conserved aromatic residues of σ region 2.3 (Figures 1C and 1E) (Juang and Helmann, 1994). The role, if any, of sequence-specific recognition of the duplex −10 element in RP_c is less clear, due to the transient nature of this intermediate. Current thinking posits that the −10 element (or at least its upstream part) may be recognized sequence specifically in double-stranded DNA (dsDNA) form (i.e., in RP_c) by residues of σ region 2.4, whereas upon strand separation and RP_o formation, residues of σ region 2.3 recognize the nt-strand bases of the −10 element (reviewed in deHaseth et al., 1998; Helmann and deHaseth, 1999; Hook-Barnard and Hinton, 2007). In contrast to this view, our structural modeling suggests that sequence-specific interactions between σ_2 and the duplex −10 element are unlikely to form prior to strand separation beginning at $\mathbf{A}_{-11}$: we hypothesize that recognition of the −10 element sequence only occurs when strand separation is initiated, as the $\mathbf{A}_{-11}$ and $\mathbf{T}_{-7}$ bases are captured in their σ_2 pockets (Figures 2, 3, and 4).

To test our hypothesis, we investigated the binding of dsDNA containing −10 element sequences to *E. coli* RNAP holoenzyme compared with DNA lacking the −10 element (anti[−10] DNA) under conditions favoring RP_c (4°C). To monitor DNA binding, we employed a recently reported "RNAP beacon assay" (Mekler et al., 2011), which takes advantage of the sensitivity of light emission from a fluorophore attached on the σ surface near the cluster of aromatic residues implicated in −10 element recognition. In free RNAP, the probe fluorescence is quenched due to photoinduced electron transfer from the cluster. Upon

binding of −10 element DNA, the contacts between the aromatic residues and the fluorophore become disrupted, resulting in increased fluorescence signal. The assay is ideal for our purposes because it reports on specific σ_2/−10 element interactions while being "blind" to nonspecific protein/DNA binding elsewhere on the holoenzyme that can mask weak, specific interactions in conventional binding assays.

To focus on RNAP/−10 element interactions and avoid the contribution of other promoter elements to binding affinity, we chose a dsDNA fragment (−22 to +4) based on the *lac*UV5 promoter (Figure 5), for which a stable RP_c has been reported at low temperatures (Spassky et al., 1985). We observed specific binding of the dsDNA fragment ($K_D = 1.0 \pm 0.4$ μM at 4°C; Figures 5A and 5D). A single-base substitution at the −11 position (A to G) resulted in a significant drop of affinity ($K_D = 33.7 \pm 14.9$ μM), almost to the level of the anti(−10) sequence ($K_D = 48.9 \pm 21.1$ μM; Figure 5A). In the two extremes, the observed specific interaction could be between RNAP holoenzyme and the fully duplex DNA fragment (as postulated in RP_c), or alternatively, the −10 element may be bound to RNAP holoenzyme with the $\mathbf{A}_{-11}$ and $\mathbf{T}_{-7}$ bases in their respective σ_2 pockets (Figure 1E). To distinguish between these two scenarios, we introduced modified bases at the −11 and −7 positions of the −10 element duplex designed to prevent binding of the bases in their σ_2 pockets but preserve the recognition surfaces of the dsDNA and measured their effects on binding.

According to available RP_c models (Murakami et al., 2002; Shultzaberger et al., 2007), the $(A/T)_{-11}$ base pair of the duplex −10 element is exposed to σ_2 via its major groove (Figure 6). With 2,6-diaminopurine (diAP) in place of $\mathbf{A}_{-11}$, the major-groove profile and overall geometry of the −11 base pair remain intact (Figure 5D) (Cheong et al., 1988), but binding of the base in its σ_2 pocket (Figures 2B and 3) is compromised due to steric clash of the exocyclic amine at the 2 position (Figure S2B) (Lee et al., 2004). The $diAP_{-11}$ incorporated into the duplex −10 element fragment caused a 14-fold decrease of binding (Figures 5B and 5D). We presumed that the residual binding of $(diAP/T)_{-11}$ dsDNA (compared to anti[−10]) was due to recognition of the −10-like sequence on the bottom strand, where 4 bases out of 6 match the consensus (Figure 5D). Neutralizing this second −10 element with a 2-aminopurine (2AP) modification in the bottom strand opposite $\mathbf{T}_{-7}$ resulted in a loss of binding nearly to the level of the anti(−10) DNA (Figures 5B and 5D). The $(T/2AP)_{-7}$ modification by itself does not affect the recognition of the nt-strand −10 element (Figure 5C). Introduction of diAP into each of three possible positions of a −35 element-containing duplex promoter fragment (−41 to −12) had no effect on binding (Figure S2C), ruling out possible effects of diAP on DNA helix geometry that could affect the putative −10 element dsDNA mode of binding.

Introducing modified bases into the t-strand would not be expected to affect the binding of the nt-strand $\mathbf{A}_{-11}$ in its σ_2 pocket. Indeed, 5-methyl isocytosine (MeiC) or 3-nitropyrrole (3-NP) opposite $\mathbf{A}_{-11}$ alters the minor (MeiC, 3-NP) and major (3-NP) grooves of the base pair, but these modifications have no significant effect on DNA binding (Figures 5B and 5D).

In RP_c, the $(T/A)_{-7}$ bp is expected to face σ_2 via its minor groove (Murakami et al., 2002; Shultzaberger et al., 2007) (Figure 6). Replacing the $(T/A)_{-7}$ bp with C/H (H, hypoxanthine) preserves the disposition of functional groups within the minor groove but prevents binding of the −7 nt-base (Figures 5C and S2A). This modification resulted in a 30-fold increase in the K_D (Figures 5C and 5D). Introducing 2AP or even a universal base (5-nitroindole; 5-NI) in the t-strand opposite $\mathbf{T}_{-7}$ had no effect on dsDNA fragment binding even though these modifications significantly alter both major and minor groove profiles of the dsDNA (Figures 5C and 5D).

Finally, $(H/C)_{-11}$ and $(2\text{-sT/A})_{-7}$ (2-sT, 2-thiothymidine) modifications address the unlikely cases where, in RP_c, the $(A/T)_{-11}$ bp faces σ_2 from its minor groove and the $(T/A)_{-7}$ bp faces σ_2 from its major groove (Figure 5D). Again, even though these alterations preserve the respective dsDNA grooves that could be facing the σ surface, these modifications to the nt-stand $\mathbf{A}_{-11}$ and $\mathbf{T}_{-7}$ bases compromise the fit in their σ_2 pockets and result in a decrease of binding affinity (Figure 5D).

In summary, we find that modifications to the $\mathbf{A}_{-11}$ or $\mathbf{T}_{-7}$ bases of the nt-strand of the −10 element expected to disrupt ssDNA binding (Figure 1E) compromise binding of the dsDNA fragment (marked red in Figure 5D). At the same time, dramatic alterations to the major and minor groove structures of the dsDNA do not significantly affect binding, as long as the $\mathbf{A}_{-11}$ or $\mathbf{T}_{-7}$ bases remain intact (marked green in Figure 5D). In combination, these results can only be explained if the critical nt-strand $\mathbf{A}_{-11}$ and $\mathbf{T}_{-7}$ bases are bound by σ_2 in the single-stranded state and not in the context of fully closed dsDNA. We conclude that the specific binding observed for the unmodified duplex −10 element fragment is due to recognition of the $\mathbf{A}_{-11}$ and $\mathbf{T}_{-7}$ bases in their σ_2 pockets, and that within the limits of detection of our assay, sequence-specific recognition of the duplex −10 element does not occur.

DISCUSSION

The −10 element was discovered more than three decades ago (Pribnow, 1975). Nevertheless, the molecular details of its recognition by the RNAP holoenzyme have, until now, been unknown. The crystal structures presented here reveal a high-resolution view of the sequence-specific interactions between the bacterial RNAP promoter-specificity σ factor and the nt-strand of the −10 element. These interactions are critical for nucleation of melting and stabilization of the initial strand-separated state that ultimately allow the formation of the transcription bubble, providing the RNAP active site access to the DNA t-strand for coding of the transcript sequence.

Base-specific interactions between σ and the ssDNA quantitatively reflect the conservation at each position of the −10 element (Figure 1B). The observed ssDNA/σ_2 interactions are completely consistent with previous footprinting and chemical probing experiments within the −10 element (Johnsrud, 1978; Siebenlist et al., 1980) and explain the strict requirements on the base at the −11 position (Figure 3C) (Lee et al., 2004; Lim et al., 2001; Matlock and Heyduk, 2000).

All of the protein amino acid side chains that interact with the ssDNA are highly conserved (most of them invariant; Figure 1C). Many σ_2 residues implicated previously in promoter binding (R237, K241 [Tomsic et al., 2001]; K249 [Waldburger and

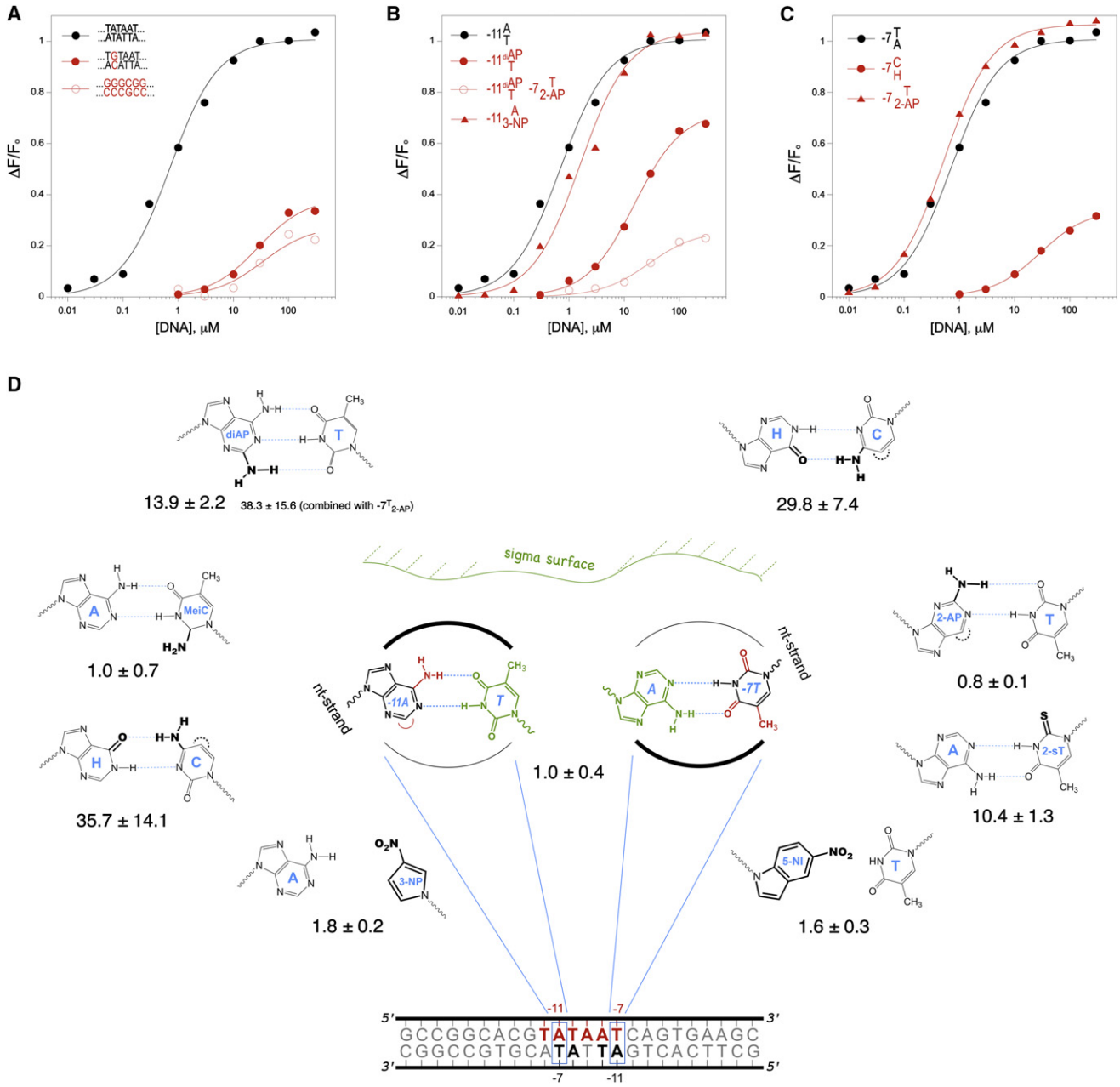

Figure 5. Interactions of −10 Element dsDNA with RNAP Holoenzyme; See also Figure S2

(A–C) Binding isotherms for dsDNA fragments to RNAP holoenzyme. Shown are representative curves for the binding of σ[211-Cys-Alexa555]-holoenzyme at increasing dsDNA concentration as measured using the RNAP beacon assay (Experimental Procedures) (Mekler et al., 2011). Sequences/modifications are shown above binding curves on the left.

(D) Summary of DNA modifications studied. Schematic showing orientation of the $(A/T)_{-11}$ and $(T/A)_{-7}$ base pairs in relation to the DNA-binding surface of σ_2. Modifications decreasing the binding of the dsDNA promoter fragment (sequence shown at the bottom) are highlighted in red, neutral modifications are green. The −10 element sequence on the nt-stand is shown in red, −10-like element sequence on the opposite strand is shown in bold. Modified base pairs investigated in this study are shown on the left (for $(A/T)_{-11}$ base pair) and right (for $(T/A)_{-7}$ base pair). Changes from the canonical A/T base pair are highlighted in bold. Measured K_D values (μM) are shown underneath each modification. For each modification, K_D measurements were independently repeated two or three times, and averages were calculated. The experimental variation among replicate measurements usually did not exceed 20% of the average value.

Susskind, 1994]; R259 [Fenton et al., 2000]) or promoter melting (F248, Y253, W256, W257 [Juang and Helmann, 1994; Schroeder et al., 2009]; T252 [Schroeder et al., 2008]) are seen to play important roles in the complex. The conserved aromatic residues of σ region 2.3 were presumed to fulfill their promoter-melting role through stacking interactions with the −10 element

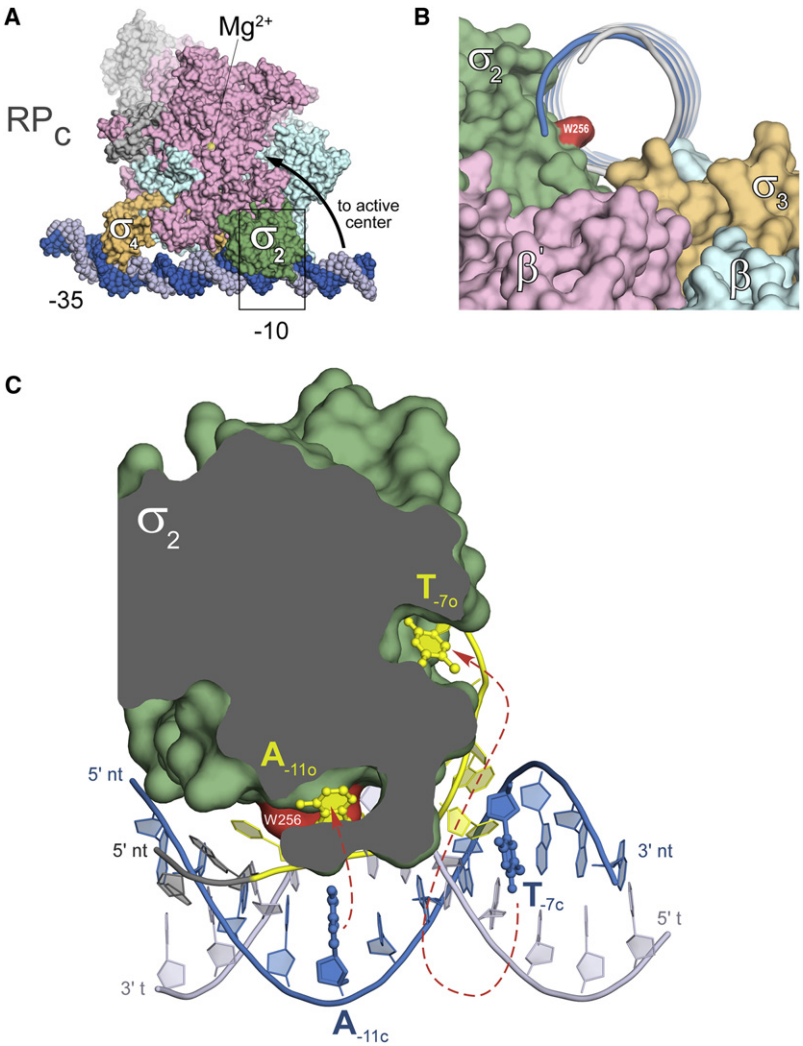

Figure 6. Structural Model of −10 Element Recognition

(A) A model of RP_c (adapted from Murakami et al., 2002). *Taq* RNAP holoenzyme is shown as a molecular surface (α subunits, ω, gray; β, cyan; β', pink; σ is orange, except σ_2, which is green). The position of the RNAP active site Mg^{2+} ion is illustrated by a yellow sphere (viewed through the β' subunit). The thick black arrow indicates how the dsDNA downstream of the −10 element must move to enter the RNAP active-site channel. The DNA t-strand is colored gray, the nt-strand blue, with the −35 and −10 elements labeled. The boxed region is shown in more detail in (C).

(B) View of the RP_c model down the DNA helix axis (the DNA is shown as a P-backbone worm) with the downstream direction into the page. The DNA double-helix sits in a shallow trough formed by σ_2, σ_3, and β. W256 of σ_2, which protrudes into the DNA helix at the position of the −11 bp, is colored red.

(C) Magnified view showing σ_2, the dsDNA of the RP_c model (with $\mathbf{A}_{-11c}$ and $\mathbf{T}_{-7c}$ highlighted), and the bound nt-strand ssDNA representing RP_o (yellow). The σ_2 is partially cut-away to reveal the $\mathbf{A}_{-11o}$ and $\mathbf{T}_{-7o}$ binding pockets. The wedge residue W256 is highlighted in red. Red arrows connect $\mathbf{A}_{-11}$ and $\mathbf{T}_{-7}$ of RP_c with the same nucleotides of RP_o.

nucleotide bases (Helmann and Chamberlin, 1988), but only Y253 participates in the complex in this way by stacking on the flipped-out $\mathbf{A}_{-11}$ (Figures 2B and 3B) (Schroeder et al., 2009). Many residues not previously implicated in promoter binding appear to make critical interactions (for instance, L108, N206, R208, L209, Figure 4; R246, Figures 2 and 3B).

Role of −10 Element in the Nucleation of Promoter Melting

During promoter opening, RNAP unwinds about 1.3 turns of the dsDNA (from −11 to +3) without any external energy input (such as ATP hydrolysis), utilizing instead the binding free energy of interactions with promoter DNA. Promoter melting, therefore, is driven by RNAP affinity toward the "final state," i.e., the conformation of promoter DNA existing in RP_o. The crucial role of σ_2 in the nucleation of promoter opening is to provide favorable interactions with the melted −10 element DNA. Specific recognition of the dsDNA in the region to be melted would stabilize the closed DNA and would therefore be unfavorable for melting. Indeed, our structural modeling and biochemical data argue that −10 element sequence readout is coupled with the nucleation of strand separation.

We show here that even under conditions that favor RP_c (4°C), RNAP specifically recognizes only the melted state of the −10 element with nt-strand bases at −11 and −7 flipped out of the DNA base stack (Figure 5). Complexes formed between RNAP holoenzyme and duplex −10 element DNA at 4°C appear "closed" on the basis of nonreactivity toward MnO_4^- oxidation, a technique used to reveal unstacked/solvent-exposed T bases (data not shown; Niedziela-Majka and Heyduk, 2005). This suggests that in intermediate complexes observed on various promoters at low temperatures, strand separation may be initiated, but the T bases may remain stacked and/or protected by protein contacts and thus be nonreactive to MnO_4^- (Davis et al., 2007). Proceeding from this initial recognition state to the final stable transcription-competent RP_o requires a combination of additional factors, such as auxiliary promoter elements, negative supercoiling, or elevated temperatures.

Relationship to Promoter-Melting Mechanisms in Higher Organisms

The cellular RNAPs from all three domains of life are conserved in sequence, structure, and catalytic mechanism (Lane and Darst, 2010), but initiation scenarios are distinct. Eukaryotic RNAPs (I, II, and III) employ an arsenal of general transcription factors (GTFs) that assemble at promoter elements located upstream and downstream of the TSS (Roeder, 2005), preparing a platform for RNAP recruitment. In contrast, bacterial σ is unable to

recognize promoters on its own and functions, therefore, as a dissociable RNAP subunit.

Some scattered functional analogies between GTFs and σ seem to have resulted from convergent evolution. In the case of RNAP II (the best understood eukaryotic initiation system), the TATA-box (reviewed in Nikolov and Burley, 1997), despite striking sequence similarity, is not analogous to the −10 element in function nor in recognition mechanism. The fact that it is recognized in double-stranded form (Nikolov and Burley, 1997) and is located ~20 base pairs upstream of the origin of melting and ~30 base pairs upstream of the TSS make its role more similar to the bacterial −35 element. Continuing this analogy, the GTFs that assemble around the TATA-box (TBP and elements of TFIIB) play roles analogous to bacterial σ_4. A "finger" domain of TFIIB, a possible functional counterpart of bacterial σ_2, may be responsible for stabilization of melting and TSS selection, but the mechanism of its action remains unknown (Bushnell et al., 2004; Kostrewa et al., 2009; Liu et al., 2010).

The hallmark of our model for the role of σ_2 in −10 element recognition, melting, and TSS selection is that these steps are spatially and temporally coupled. By contrast, in eukaryotic initiation, these steps appear to be independent and follow complex, stepwise mechanisms allowing for multiple layers of regulation (Kostrewa et al., 2009; Liu et al., 2010; Roeder, 2005).

Dynamics of Promoter-Melting Nucleation

Many studies have suggested that strand separation initiates at −11 and then propagates downstream (Chen and Helmann, 1995; Heyduk et al., 2006; Lim et al., 2001). Recent computational modeling found that initiation of strand separation at −11 resulted in efficient kinetics of RP_o formation, whereas bubble initiation elsewhere yielded inefficient trajectories (Chen et al., 2010). A structural basis for these observations, as well as other insights into the initiation of transcription bubble formation, is provided by a comparison of the structure presented here (representing RP_o) with a simple RP_c model (adapted from Murakami et al., 2002; Figure 6).

The flipping of the $\mathbf{A}_{-11}$ base is key to initiating the strand-separation process (Heyduk et al., 2006; Lim et al., 2001). We noted that the absolutely conserved σ_2-W256 resides on the downstream face of $\mathbf{T}_{-12}$, precisely where $\mathbf{A}_{-11}$ would be if the dsDNA continued downstream to −11 (Figures 2B and 3A). In the holoenzyme structure, the bulky W256 side chain protrudes from the bottom of a shallow, electrostatically basic trough with dimensions to accommodate dsDNA and formed by surfaces of σ_2, σ_3, and the β subunit (Figure 6B). This prominent position of W256 at the at the −11 register of RP_c suggests that when dsDNA is loaded in this trough, directed by electrostatic interactions with the DNA backbone, the W256 side chain may act as a "wedge" for disrupting the $(A/T)_{-11}$ base pair, initiating $\mathbf{A}_{-11}$ flipping (active mechanism) or stabilizing the conformation of the spontaneously flipped $\mathbf{A}_{-11}$ base (passive mechanism). These structural considerations corroborate initial suggestions for a crucial role of W256 in $\mathbf{A}_{-11}$ flipping based on the observation that the substitution of W433A in *E. coli* σ^{70} (corresponding to *Taq* σ^A W256) had no effect on single-stranded −10 element binding but dramatically slowed the rate of double-stranded promoter DNA opening (Tomsic et al., 2001).

The proposed role of W256 in $\mathbf{A}_{-11}$ recognition invokes a mechanistic analogy to base flipping by other DNA-binding proteins that also use a wedge residue (often an aromatic side chain) to invade the DNA double-helix and fill the space vacated by the flipped base, stabilizing its extrahelical conformation. Examples of such base-flipping proteins are numerous and include base excision repair proteins (Lau et al., 1998; Yang et al., 2009) or Tn5 transposase (that also uses a W residue as a wedge for base flipping; Davies et al., 2000).

Although the −12 position likely remains base-paired even in RP_o, structural considerations indicate that the $(T/A)_{-12}$ base pair can only be recognized after the $\mathbf{A}_{-11}$ is removed from the double-helix (Figure 3A), suggesting that sequence-specific readout of the first two upstream positions of the −10 element occur at the $\mathbf{A}_{-11}$ flipping step. Subsequently, DNA helix untwisting continues downstream, driven by σ_2 interaction with the sugar-phosphate backbone of the central nonconserved part of the −10 hexamer $\mathbf{T}_{-10}\mathbf{A}_{-9}\mathbf{A}_{-8}$ (Figure 4). This brings $\mathbf{T}_{-7}$ closer to its pocket on σ_2. The recognition of $\mathbf{T}_{-7}$, and the discriminator element further downstream, may serve as "check points" along the pathway of propagation of melting toward the TSS (Figure 6) (Matlock and Heyduk, 2000).

Formation of the final RP_o requires interactions between RNAP core subunits in the active-site cleft with promoter DNA downstream of the −10 element (Saecker et al., 2002), but σ_2/−10 element recognition is sufficient for the initial strand separation (Young et al., 2004). The key bases recognized by σ_2, $\mathbf{A}_{-11}$ and $\mathbf{T}_{-7}$, are roughly 180° apart on the DNA helical axis (Figure 6C), therefore their binding in the pockets, as observed in our crystal structures, results in untwisting of the double-helix by about half a turn. This initial untwisting also necessitates a sharp bend in the nt-stand (90° kink between positions −11 and −10, as observed in our structure; Figure 1E), which is likely responsible for positioning the downstream dsDNA in the RNAP active-site channel (Saecker et al., 2002; Figure 6A).

Whether σ actively disrupts the $(A/T)_{-11}$ base pair or passively captures transiently exposed base(s) remains to be established—the two pathways are not mutually exclusive. Double-stranded DNA is thermodynamically stable but kinetically labile; individual base pairs have an average lifetime on the order of milliseconds (Guéron and Leroy, 1995) and are in equilibrium with flipped-out bases. Particularly unstable are "TA" steps (as found in the −10 element) due to the relatively weak stacking interactions (Protozanova et al., 2004). Indeed, solution studies of bacterial promoters have shown that the −10 element has an altered structure, even in the absence of proteins (Drew et al., 1985; Spassky et al., 1988). The detailed mechanism of initial base flipping in the −10 element poses intriguing questions for future research.

EXPERIMENTAL PROCEDURES

Full details of experimental procedures are presented in the Extended Experimental Procedures online.

Protein and Nucleic Acid Preparation

The *Taq* σ^A_{2-3} fragment was subcloned into a pET28a-derived expression vector, transformed into *E. coli* BL21(DE3) cells, overexpressed, and purified using standard methods. The purified σ^A_{2-3} was concentrated to 40 mg/ml

by centrifugal filtration (VivaScience) in 10 mM Tris, pH 8.0, 150 mM KCl, 0.1 mM EDTA, then flash frozen and stored at −80°C. The PAGE-purified oligonucleotides (Oligos Etc.) were dissolved in water to a concentration of 3 mM prior to use.

Crystallization

The ssDNA/σ^{A}_{2-3} complex was prepared on ice (molar ratio 2:1) at a final protein concentration of 10 mg/ml. Crystals were grown at 22°C using hanging-drop vapor diffusion by mixing equal volumes of the complex and a reservoir solution of 100 mM Tris, pH 8.5, 5% (w/v) PEG 8000, 20% (w/v) PEG 300, 10% (v/v) glycerol, and 0.15% (w/v) mellitic acid. For data collection, crystals were flash-frozen in liquid nitrogen directly from the mother liquor.

Structure Determination

Diffraction data were collected at the Advanced Photon Source (Argonne National Laboratory) beamline NE-CAT 24 ID-E and at the National Synchrotron Light Source (Brookhaven National Laboratory) beamline X29. The structure was solved by molecular replacement. Iterative rounds of model building and refinement yielded the final models (Table S1).

RNAP Beacon Assay

The RNAP beacon assay was performed essentially as described (Mekler et al., 2011), except that Alexa 555 fluorophore was used.

ACCESSION NUMBERS

The X-ray crystallographic coordinates and structure factor files have been deposited in the Protein Data Bank with accession IDs 3UGO (TGTACAATGGG oligo) and 3UGP (TGTATAATGGG oligo).

SUPPLEMENTAL INFORMATION

Supplemental Information includes two figures and one table and can be found with this article online at doi:10.1016/j.cell.2011.10.041.

ACKNOWLEDGMENTS

We thank R. Shultzaberger and T. Schneider for providing Figure 1B and for helpful discussions; P. deHaseth, A. Kulbachinskiy, S. Malik, A. Mustaev, R. Saecker, and M. Schapira for helpful discussions; K.R. Rajashankar and F. Murphy at APS NE-CAT beamline 24ID-E and W. Shi at NSLS beamline X29 for support with synchrotron data collection; and R. MacKinnon for use of the fluorescence plate-reader. A.F. is a Merck Postdoctoral Fellow at The Rockefeller University. This work was based, in part, on research conducted at the APS and the NSLS, supported by the US Department of Energy, Office of Basic Energy Sciences. The NE-CAT beamlines at the APS are supported by award RR-15301 from the NCRR at the NIH. This work was supported by NIH RO1 GM053759 to S.A.D.

Received: August 24, 2011
Revised: October 4, 2011
Accepted: October 6, 2011
Published online: December 1, 2011

REFERENCES

Bushnell, S.A., Westover, K.D., Davis, R.E., and Kornberg, R.D. (2004). Structural basis of transcription: an rNA polymerase II-TFIIB cocrystal at 4.5 Angstroms. Science *303*, 983–988.

Campbell, E.A., Muzzin, O., Chlenov, M., Sun, J.L., Olson, C.A., Weinman, O., Trester-Zedlitz, M.L., and Darst, S.A. (2002). Structure of the bacterial RNA polymerase promoter specificity sigma subunit. Mol. Cell *9*, 527–539.

Chen, J., Darst, S.A., and Thirumalai, D. (2010). Promoter melting triggered by bacterial RNA polymerase occurs in three steps. Proc. Natl. Acad. Sci. USA *107*, 12523–12528.

Chen, Y.F., and Helmann, J.D. (1995). The *Bacillus subtilis* flagellar regulatory protein σ^{D}: overproduction, domain analysis and DNA-binding properties. J. Mol. Biol. *249*, 743–753.

Cheong, C., Tinoco, I.J., Jr., and Chollet, A. (1988). Thermodynamic studies of base pairing involving 2,6-diaminopurine. Nucleic Acids Res. *16*, 5115–5122.

Davies, D.R., Goryshin, I.Y., Reznikoff, W.S., and Rayment, I. (2000). Three-dimensional structure of the Tn5 synaptic complex transposition intermediate. Science *289*, 77–85.

Davis, C.A., Bingman, C.A., Landick, R., Record, M.T.J., Jr., and Saecker, R.M. (2007). Real-time footprinting of DNA in the first kinetically significant intermediate in open complex formation by *Escherichia coli* RNA polymerase. Proc. Natl. Acad. Sci. USA *104*, 7833–7838.

deHaseth, P.L., and Helmann, J.D. (1995). Open complex formation by *Escherichia coli* RNA polymerase: the mechanism of polymerase-induced strand separation of double helical DNA. Mol. Microbiol. *16*, 817–824.

deHaseth, P.L., Zupancic, M.L., and Record, M.T.J., Jr. (1998). RNA polymerase-promoter interactions: the comings and goings of RNA polymerase. J. Bacteriol. *180*, 3019–3025.

Drew, H.R., Weeks, J.R., and Travers, A.A. (1985). Negative supercoiling induces spontaneous unwinding of a bacterial promoter. EMBO J. *4*, 1025–1032.

Feklistov, A., Barinova, N., Sevostyanova, A., Heyduk, E., Bass, I., Vvedenskaya, I., Kuznedelov, K., Merkiene, E., Stavrovskaya, E., Klimasauskas, S., et al. (2006). A basal promoter element recognized by free RNA polymerase σ subunit determines promoter recognition by RNA polymerase holoenzyme. Mol. Cell *23*, 97–107.

Fenton, M.S., and Gralla, J.D. (2001). Function of the bacterial TATAAT -10 element as single-stranded DNA during RNA polymerase isomerization. Proc. Natl. Acad. Sci. USA *98*, 9020–9025.

Fenton, M.S., and Gralla, J.D. (2003a). Effect of DNA bases and backbone on sigma70 holoenzyme binding and isomerization using fork junction probes. Nucleic Acids Res. *31*, 2745–2750.

Fenton, M.S., and Gralla, J.D. (2003b). Roles for inhibitory interactions in the use of the -10 promoter element by sigma 70 holoenzyme. J. Biol. Chem. *278*, 39669–39674.

Fenton, M.S., Lee, S.J., and Gralla, J.D. (2000). *Escherichia coli* promoter opening and -10 recognition: mutational analysis of sigma70. EMBO J. *19*, 1130–1137.

Gruber, T.M., and Bryant, D.A. (1997). Molecular systematic studies of eubacteria, using sigma70-type sigma factors of group 1 and group 2. J. Bacteriol. *179*, 1734–1747.

Guéron, M., and Leroy, J.L. (1995). Studies of base pair kinetics by NMR measurement of proton exchange. Methods Enzymol. *261*, 383–413.

Guo, Y., and Gralla, J.D. (1998). Promoter opening via a DNA fork junction binding activity. Proc. Natl. Acad. Sci. USA *95*, 11655–11660.

Haugen, S.P., Berkmen, M.B., Ross, W., Gaal, T., Ward, C., and Gourse, R.L. (2006). rRNA promoter regulation by nonoptimal binding of σ region 1.2: an additional recognition element for RNA polymerase. Cell *125*, 1069–1082.

Haugen, S.P., Ross, W., Manrique, M., and Gourse, R.L. (2008). Fine structure of the promoter-sigma region 1.2 interaction. Proc. Natl. Acad. Sci. USA *105*, 3292–3297.

Helmann, J.D., and Chamberlin, M.J. (1988). Structure and function of bacterial sigma factors. Annu. Rev. Biochem. *57*, 839–872.

Helmann, J.D., and deHaseth, P.L. (1999). Protein-nucleic acid interactions during open complex formation investigated by systematic alteration of the protein and DNA binding partners. Biochemistry *38*, 5959–5967.

Heyduk, E., Kuznedelov, K., Severinov, K., and Heyduk, T. (2006). A consensus adenine at position -11 of the nontemplate strand of bacterial promoter is important for nucleation of promoter melting. J. Biol. Chem. *281*, 12362–12369.

Hoffman, M.M., Khrapov, M.A., Cox, J.C., Yao, J., Tong, L., and Ellington, A.D. (2004). AANT: the amino acid-nucleotide interaction database. Nucleic Acids Res. *32*(Database issue), D174–D181.

Hook-Barnard, I.G., and Hinton, D.M. (2007). Transcription initiation by mix and match elements: flexibility for polymerase binding to bacterial promoters. Gene Regul Syst Bio *1*, 275–293.

Johnsrud, L. (1978). Contacts between *Escherichia coli* RNA polymerase and a lac operon promoter. Proc. Natl. Acad. Sci. USA *75*, 5314–5318.

Juang, Y.-L., and Helmann, J.D. (1994). A promoter melting region in the primary sigma factor of Bacillus subtilis. Identification of functionally important aromatic amino acids. J. Mol. Biol. *235*, 1470–1488.

Kenney, T.J., York, K., Youngman, P., and Moran, C.P.J., Jr. (1989). Genetic evidence that RNA polymerase associated with σ^{A} factor uses a sporulation-specific promoter in *Bacillus subtilis*. Proc. Natl. Acad. Sci. USA *86*, 9109–9113.

Kostrewa, D., Zeller, M.E., Armache, K.-J., Seizl, M., Leike, K., Thomm, M., and Cramer, P. (2009). RNA polymerase II-TFIIB structure and mechanism of transcription initiation. Nature *462*, 323–330.

Lane, W.J., and Darst, S.A. (2010). Molecular evolution of multisubunit RNA polymerases: sequence analysis. J. Mol. Biol. *395*, 671–685.

Lau, A.Y., Schärer, O.D., Samson, L., Verdine, G.L., and Ellenberger, T. (1998). Crystal structure of a human alkylbase-DNA repair enzyme complexed to DNA: mechanisms for nucleotide flipping and base excision. Cell *95*, 249–258.

Lee, H.J., Lim, H.M., and Adhya, S. (2004). An unsubstituted C2 hydrogen of adenine is critical and sufficient at the -11 position of a promoter to signal base pair deformation. J. Biol. Chem. *279*, 16899–16902.

Lim, H.M., Lee, H.J., Roy, S., and Adhya, S. (2001). A "master" in base unpairing during isomerization of a promoter upon RNA polymerase binding. Proc. Natl. Acad. Sci. USA *98*, 14849–14852.

Liu, X., Bushnell, D.A., Wang, D., Calero, G., and Kornberg, R.D. (2010). Structure of an RNA polymerase II-TFIIB complex and the transcription initiation mechanism. Science *327*, 206–209.

Lonetto, M., Gribskov, M., and Gross, C.A. (1992). The σ^{70} family: sequence conservation and evolutionary relationships. J. Bacteriol. *174*, 3843–3849.

Luscombe, N.M., Laskowski, R.A., and Thornton, J.M. (2001). Amino acid-base interactions: a three-dimensional analysis of protein-DNA interactions at an atomic level. Nucleic Acids Res. *29*, 2860–2874.

Marr, M.T., and Roberts, J.W. (1997). Promoter recognition as measured by binding of polymerase to nontemplate strand oligonucleotide. Science *276*, 1258–1260.

Matlock, D.L., and Heyduk, T. (2000). Sequence determinants for the recognition of the fork junction DNA containing the -10 region of promoter DNA by *E. coli* RNA polymerase. Biochemistry *39*, 12274–12283.

Mekler, V., Pavlova, O., and Severinov, K. (2011). Interaction of *Escherichia coli* RNA polymerase σ^{70} subunit with promoter elements in the context of free σ^{70}, RNA polymerase holoenzyme, and the β'-σ^{70} complex. J. Biol. Chem. *286*, 270–279.

Moyle, H., Waldburger, C., and Susskind, M.M. (1991). Hierarchies of base pair preferences in the P22 ant promoter. J. Bacteriol. *173*, 1944–1950.

Murakami, K.S., and Darst, S.A. (2003). Bacterial RNA polymerases: the wholo story. Curr. Opin. Struct. Biol. *13*, 31–39.

Murakami, K.S., Masuda, S., Campbell, E.A., Muzzin, O., and Darst, S.A. (2002). Structural basis of transcription initiation: an RNA polymerase holoenzyme-DNA complex. Science *296*, 1285–1290.

Niedziela-Majka, A., and Heyduk, T. (2005). Escherichia coli RNA polymerase contacts outside the -10 promoter element are not essential for promoter melting. J. Biol. Chem. *280*, 38219–38227.

Nikolov, D.B., and Burley, S.K. (1997). RNA polymerase II transcription initiation: a structural view. Proc. Natl. Acad. Sci. USA *94*, 15–22.

Pribnow, D. (1975). Nucleotide sequence of an RNA polymerase binding site at an early T7 promoter. Proc. Natl. Acad. Sci. USA *72*, 784–788.

Protozanova, E., Yakovchuk, P., and Frank-Kamenetskii, M.D. (2004). Stacked-unstacked equilibrium at the nick site of DNA. J. Mol. Biol. *342*, 775–785.

Ring, B.Z., Yarnell, W.S., and Roberts, J.W. (1996). Function of E. coli RNA polymerase sigma factor sigma 70 in promoter-proximal pausing. Cell *86*, 485–493.

Roberts, C.W., and Roberts, J.W. (1996). Base-specific recognition of the nontemplate strand of promoter DNA by *E. coli* RNA polymerase. Cell *86*, 495–501.

Roeder, R.G. (2005). Transcriptional regulation and the role of diverse coactivators in animal cells. FEBS Lett. *579*, 909–915.

Saecker, R.M., Tsodikov, O.V., McQuade, K.L., Schlax, P.E.J., Jr., Capp, M.W., and Record, M.T.J., Jr. (2002). Kinetic studies and structural models of the association of *E. coli* sigma(70) RNA polymerase with the lambdaP(R) promoter: large scale conformational changes in forming the kinetically significant intermediates. J. Mol. Biol. *319*, 649–671.

Savinkova, L.K., Baranova, L.V., Knorre, V.L., and Salganik, R.I. (1988). [Binding of RNA-polymerase from Escherichia coli with oligodeoxyribonucleotides homologous to transcribed and non-transcribed DNA stands in the "-10"-promoter region of bacterial genes]. Mol. Biol. (Mosk.) *22*, 807–812.

Schneider, T.D. (2001). Strong minor groove base conservation in sequence logos implies DNA distortion or base flipping during replication and transcription initiation. Nucleic Acids Res. *29*, 4881–4891.

Schneider, T.D., and Stephens, R.M. (1990). Sequence logos: a new way to display consensus sequences. Nucleic Acids Res. *18*, 6097–6100.

Schroeder, L.A., Choi, A.-J., and DeHaseth, P.L. (2007). The -11A of promoter DNA and two conserved amino acids in the melting region of sigma70 both directly affect the rate limiting step in formation of the stable RNA polymerase-promoter complex, but they do not necessarily interact. Nucleic Acids Res. *35*, 4141–4153.

Schroeder, L.A., Karpen, M.E., and deHaseth, P.L. (2008). Threonine 429 of *Escherichia coli* sigma 70 is a key participant in promoter DNA melting by RNA polymerase. J. Mol. Biol. *376*, 153–165.

Schroeder, L.A., Gries, T.J., Saecker, R.M., Record, M.T.J., Jr., Harris, M.E., and DeHaseth, P.L. (2009). Evidence for a tyrosine-adenine stacking interaction and for a short-lived open intermediate subsequent to initial binding of Escherichia coli RNA polymerase to promoter DNA. J. Mol. Biol. *385*, 339–349.

Severinova, E., Severinov, K., Fenyö, D., Marr, M., Brody, E.N., Roberts, J.W., Chait, B.T., and Darst, S.A. (1996). Domain organization of the *Escherichia coli* RNA polymerase σ^{70} subunit. J. Mol. Biol. *263*, 637–647.

Sevostyanova, A., Feklistov, A., Barinova, N., Heyduk, E., Bass, I., Klimasauskas, S., Heyduk, T., and Kulbachinskiy, A. (2007). Specific recognition of the -10 promoter element by the free RNA polymerase sigma subunit. J. Biol. Chem. *282*, 22033–22039.

Shultzaberger, R.K., Chen, Z., Lewis, K.A., and Schneider, T.D. (2007). Anatomy of *Escherichia coli* sigma70 promoters. Nucleic Acids Res. *35*, 771–788.

Siebenlist, U., Simpson, R.B., and Gilbert, W. (1980). *E. coli* RNA polymerase interacts homologously with two different promoters. Cell *20*, 269–281.

Spassky, A., Kirkegaard, K., and Buc, H. (1985). Changes in the DNA structure of the *lac* UV5 promoter during formation of an open complex with *Escherichia coli* RNA polymerase. Biochemistry *24*, 2723–2731.

Spassky, A., Rimsky, S., Buc, H., and Busby, S. (1988). Correlation between the conformation of *Escherichia coli* -10 hexamer sequences and promoter strength: use of orthophenanthroline cuprous complex as a structural index. EMBO J. *7*, 1871–1879.

Tomsic, M., Tsujikawa, L., Panaghie, G., Wang, Y., Azok, J., and deHaseth, P.L. (2001). Different roles for basic and aromatic amino acids in conserved region 2 of *Escherichia coli* sigma(70) in the nucleation and maintenance of the single-stranded DNA bubble in open RNA polymerase-promoter complexes. J. Biol. Chem. *276*, 31891–31896.

Tsujikawa, L., Strainic, M.G., Watrob, H., Barkley, M.D., and DeHaseth, P.L. (2002). RNA polymerase alters the mobility of an A-residue crucial to polymerase-induced melting of promoter DNA. Biochemistry *41*, 15334–15341.

Waldburger, C., Gardella, T., Wong, R., and Susskind, M.M. (1990). Changes in conserved region 2 of *Escherichia coli* sigma 70 affecting promoter recognition. J. Mol. Biol. *215*, 267–276.

Waldburger, C., and Susskind, M.M. (1994). Probing the informational content of *Escherichia coli* sigma 70 region 2.3 by combinatorial cassette mutagenesis. J. Mol. Biol. *235*, 1489–1500.

Wintjens, R., Liévin, J., Rooman, M., and Buisine, E. (2000). Contribution of cation-pi interactions to the stability of protein-DNA complexes. J. Mol. Biol. *302*, 395–410.

Yang, C.-G., Garcia, K., and He, C. (2009). Damage detection and base flipping in direct DNA alkylation repair. ChemBioChem *10*, 417–423.

Young, B.A., Anthony, L.C., Gruber, T.M., Arthur, T.M., Heyduk, E., Lu, C.Z., Sharp, M.M., Heyduk, T., Burgess, R.R., and Gross, C.A. (2001). A coiled-coil from the RNA polymerase beta' subunit allosterically induces selective nontemplate strand binding by sigma(70). Cell *105*, 935–944.

Young, B.A., Gruber, T.M., and Gross, C.A. (2004). Minimal machinery of RNA polymerase holoenzyme sufficient for promoter melting. Science *303*, 1382–1384.

Zenkin, N., Kulbachinskiy, A., Yuzenkova, Y., Mustaev, A., Bass, I., Severinov, K., and Brodolin, K. (2007). Region 1.2 of the RNA polymerase σ subunit controls recognition of the -10 promoter element. EMBO J. *26*, 955–964.